

1

CHAPTER ONE

Using Access in Forensic Investigations

FORENSIC ANALYTICS IS THE procurement and analysis of electronic data to reconstruct, detect, or otherwise support a claim of financial fraud. The main steps in forensic analytics are (a) data collection, (b) data preparation, (c) data analysis, and (d) reporting. This book casts a wider net than simply the detection of financial fraud. Using computer-based analytic methods our goal is the detection of fraud, errors, and biases where biases involve people gravitating to specific numbers or number ranges to circumvent actual or perceived internal control thresholds. These analytic methods are directed at determining the likelihood or magnitude of fraud occurring. They would be a part of a fraud deterrence cycle that would include other steps such as employment screening procedures, including background checks. The techniques described in the book rely on the analysis of data, usually transactional data, but at times, other data such as statistical data or aggregated data of some sort.

The main workhorses for the preparation and analysis of data will be Microsoft Access and Microsoft Excel (or Access and Excel, for short). Other valuable and dependable and high-quality tools for data analysis include IDEA, Minitab, and SigmaPlot for preparing high-quality complex graphs. The reporting and presentation of the results is usually done using Microsoft Word and/or Microsoft PowerPoint. These results could include images cropped from various sources (including Access and Excel). Images can be copied and pasted into Word or PowerPoint by using a software tool called Snag-It.

This chapter introduces Access and the components and features of Access that are used in a forensic analytics environment. The next two chapters do the same for Excel and PowerPoint. In summary, Access has almost everything that is needed for a forensic analytics application with reasonably sized data sets, where there is not a high

2 ■ Using Access in Forensic Investigations

requirement for high security. Forensic-related applications can be created in Access and other users with little or no knowledge of Access could use the system. The chapter reviews the Access components and features that make it useful for forensic analytics.

AN INTRODUCTION TO ACCESS

Access is Windows-based and so, fortunately, all the basic Windows operations work in Access. Your trusted mouse works just like before with right clicks, left clicks, and double clicks. Access is launched just like any other program using a shortcut or the **Start** button. Copying, moving, naming, and deleting files are done as usual. There are some differences that are mainly related to the fact that Access is a database program that expects the data tables to be continually changed and updated.

Access differs from Word and Excel in that for most users there was no migration from other products. Microsoft did an excellent job in showing people how to do task *x* in Word given that you used to do task *x* following a set of procedures using perhaps WordPerfect or Wordstar. Microsoft also showed people how to do task *y* in Excel given that you used to do task *y* using a series of steps in perhaps Quattro Pro or Lotus 1-2-3. For example, you can still enter `@sum(B1..B5)` in cell **B6** in Excel (2007) and not only will it calculate the sum correctly, but it will convert the formula to `=SUM(B1:B5)` for you. There is no help in Access geared to making you more familiar with the program, because there was not a preceding product that users were used to. This makes the logic of Access a little tricky to follow at first. With practice comes familiarity, and it will not be too long before you will prefer to use Access for those projects that are more suited to Access than to Excel.

One reason for favoring Access over Excel for forensic analytics work is that Access forces some discipline onto the data analysis project. Excel is basically a large free-form rectangle divided into smaller rectangles (called cells). In these cells you can (a) paste images, (b) enter numbers, (c) enter formulas, or (d) display a graph (called a chart in Excel). When you view a number in Excel, unless you click on the cell itself, you are never really sure if this is a data point or the result of a formula (a calculation). Excel is (unfortunately) very forgiving in that a column heading can be repeated (you can call both columns A and B, *People*), Excel does not mind if you call a column *Dollars* and immediately below the field name you enter the word *Rambo*. Excel has some built-in documenting capabilities (including the ability to Insert Comment) but most of the structure and the integrity are left up to the user. Without clear documentation it is easy for another user to have no clue as to what is happening in a complex spreadsheet, and even the original developer might have trouble figuring out what is happening if they look at a complex spreadsheet six months later. The opening screen for Access 2007 is shown in Figure 1.1.

In contrast to Access, most computer programs will at least do something once opened. For example, in PowerPoint you can immediately click on the blank slide and type a title or some text. This is not the case with Access. To get Access to start working



FIGURE 1.1 Opening Screen for Microsoft Access 2007

you either need to open an existing file or you need to create a new blank database. For a new forensic analytics project, the **New Blank Database** is the starting point. Clicking on **Blank Database** will start the series of dialog boxes creating a new Access database. The next step is shown in Figure 1.2.

Figure 1.2 shows the step needed to create an Access database named *Chapter1a.accdb* in a folder named *DataDrivenForensics*. Clicking the **Create** button will give the result in Figure 1.3.

The opening screen of the new database named *Chapter1a* is shown in Figure 1.3. *Table 1* is shown in the open objects panel and this is there so that the spot does not look empty. The table disappears once a new table is created and *Table 1* is closed. The navigation pane on the left lists all the Access objects and the details can be shortened or extended by selecting the drop down arrow and selecting **Object Type** or **All Access Objects**. The architecture of Access and the components of a database are discussed in the next section.

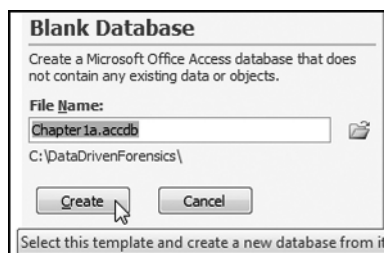


FIGURE 1.2 Creation of a New Blank Database in the *DataDrivenForensics* Folder

4 ■ Using Access in Forensic Investigations

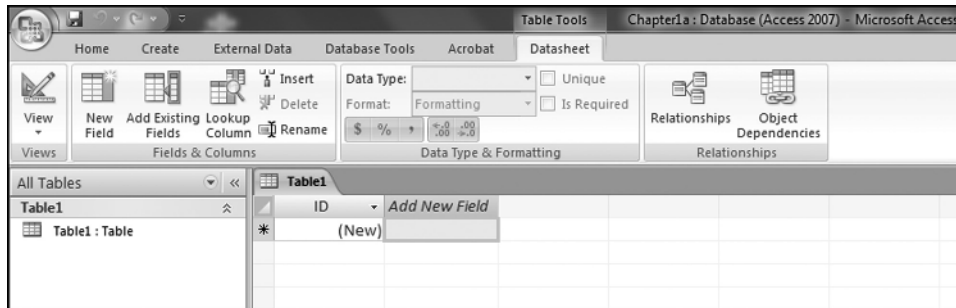


FIGURE 1.3 Opening Screen of a New Access Database Named Chapter1a

THE ARCHITECTURE OF ACCESS

The Microsoft Access homepage at <http://office.microsoft.com/en-us/access-help/> has lots of useful and reliable information on Access 2003, 2007, and 2010. The website's opening screen with Access 2007 selected is shown in Figure 1.4.

Extensive Microsoft Access information and help is available as can be seen in Figure 1.4. After selecting the appropriate version on the right (see the arrow in Figure 1.4) the site provides information and help related to using Access. A good starting place, irrespective of your Access version, is the Access Basics section in Access 2010. The *basics* are basically the same for each version except that Access 2007 and Access 2010 use the ribbon for the selection of tasks. There are also other websites with Access information and several of these are listed on the companion site for this book.

An Access database is a tool for collecting, storing, and analyzing data, and reporting information. A database consists of unprocessed data and other objects associated with collecting, editing, adding, deleting, processing, organizing, reporting on, and sharing the data. The objects listed below are of most interest from a forensic analytics perspective:



FIGURE 1.4 Microsoft Website with Access Information and Help

- **Tables.** Transaction data is stored in one or more tables. The layout of a table is the same as the layout of an Excel worksheet. Each row in the table is called a record and a record holds all the known information about one item or subject. These items or subjects could be employees, transactions, or books. The fields (columns) store similar data or facts for the various records. In a table of transactions, examples of possible fields are invoice date, invoice number, vendor number, invoice amount, and so on. In a table of census data by county examples of possible fields are county number, county name, state, area, count of people 2010, and projected count of people 2015. It is good practice to have an ID field in each table. This field is also called a primary key and holds a unique number for each record so that you can identify the record uniquely.
- **Queries.** Queries are fundamental to forensic analytics and many other Access-related tasks. Queries are often used to select a subset of records that meet certain criteria. For example, a query could retrieve all the counties in Texas with a population of less than 1,000 people. Every forensic question in Access will need a query. There are also other data-related tasks that require queries and these include appending data and updating data in tables. Queries are the workhorses of forensic analytics.
- **Reports.** Reports are used for the neat presentation of the results of the forensic analytics work. The reporting features and routines in Access allow for the creation of very neat and professional-looking reports. These reports can include conditional formatting for highlighting data. The reports can include professional-looking headings including company logos and other images. The report's footer also has many useful versatile features and capabilities. The reports can be previewed, printed on paper, viewed on a screen, exported to another program, and even converted to pdf files and sent as an attachment to an e-mail message.
- **Forms.** Forms are a user interface that can be used to enter data into tables or to edit existing data in tables. Forms can vary from being complex with command buttons and input controls to being just a basic screen with areas for data entry. Forms can also be used to neatly display the results of queries or to provide a neat way to input data. The form most often used in forensic analytics is called a switchboard. The switchboard has command buttons that can run queries or prepare reports with a single click. Switchboards allow users who are not familiar with Access to run a query or prepare a report.

Access databases can also include macros. Macros are generally time-saving objects. Macros can be used to automate tasks such as opening a report, running a query, or closing a database. The procedures for creating macros are reviewed on the Microsoft website or in any comprehensive Access book.

Access databases can also include modules that are procedures written in *Visual Basic for Applications* (VBA) that add functionality to a database. A module is a set of declarations, statements, and procedures that form a unit because they relate to one clearly defined task. Modules are flexible and we can do much more with modules than can be done by using the usual query design modes (using the design grid, SQL view, or a

6 ■ Using Access in Forensic Investigations

Wizard). Getting started with VBA requires an upfront learning curve and the good news is that all the forensic analytics tests in this book can be done without modules.

For our forensic applications we always use tables and queries. Tables hold the raw data, and queries are used to analyze the data and also to update and manipulate tables (perhaps using append queries). Reports might, or might not, be needed for neatly formatted output, and the only form that fits well with data analysis is the switchboard.

A REVIEW OF ACCESS TABLES

Tables are the starting point for any forensic analytics project. Data is stored in tables and a database can be made up of many tables. An example of a database with several tables is shown in Figure 1.5.

The database included tables for data related to a large chain of restaurants. One goal in database design is to avoid storing duplicate information (also known as *redundant data*). This reduces storage costs, the chances of data inconsistencies, and

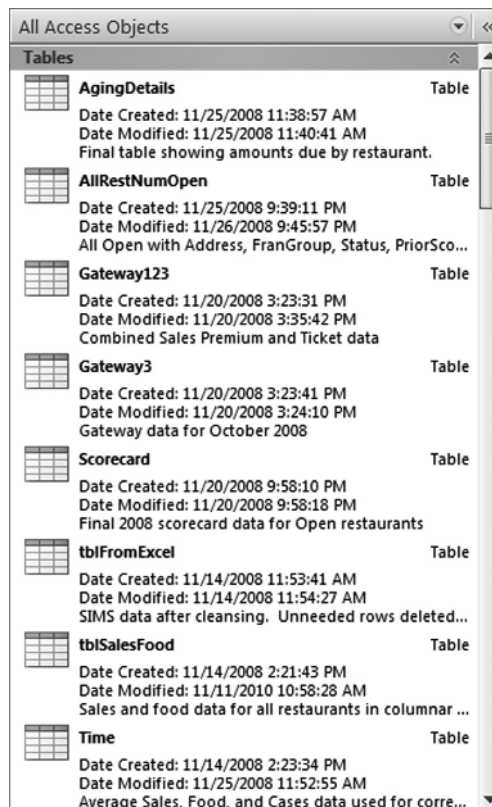


FIGURE 1.5 Access Database with Several Tables that Have Names, Descriptions, a Created Date, and a Modified Date

simplifies the task of updating records. Another principle of database design is that the database is divided into tables that each stores a relevant part of the total picture. A single table might work in some applications. Another goal is that the tables can be linked in some meaningful manner. Each restaurant in the example in Figure 1.5 has a unique restaurant number and that number (called a primary key) can be used for queries that use more than one table.

Tables are made up of records and fields. Each record contains all the information about one instance of the table subject. If the table has details about the books in a library, then each record would relate to a single book in the library. A field contains data about one aspect of the table subject. In the library example we might have a field for the book's title and another field for the acquisition date. Each record consists of field values which are also called facts. A field value might be *Lesla* or *Car* or \$19.64. There are many data types of which numeric data, dates, and text data are most applicable to forensic analytics.

For most forensic applications the data will be imported into Access from another program or from a flat file. A file with more than one million records is quite normal. The desired properties of an imported data table or of a created table are listed below:

- Each field value should contain one value only such as one date, one amount, one census count, or one first name. Text fields can use more than one word if this describes an attribute of the record, such as *New Jersey* or *Loveland Supply Company* for vendor name. In contrast, F46bl could indicate that the person is a female, 46 years old, with blue eyes, but storing all this in one field value is not good practice. The investigator would then not be able to group by *Gender* and calculate descriptive statistics, or group by *Age* and calculate descriptive statistics. The correct practice would be to have one field for each of gender, age, and eye color.
- Each field should have a distinct name. Access allows users to add a caption in the Field Properties to more fully describe the field. This caption is very useful when using databases created by other people.
- All field values should hold a value for that field only and all the field values should be of the same data type (e.g., text, or numeric, or date). A blank field value is acceptable. For example, in a table of addresses, one field might be used for the apartment or suite number and in some cases this number would not be applicable and so the field value might be blank. A blank field value is also called a *null value* for numeric data, or a *zero-length string* for text, memo, or hyperlink fields.
- The order of the records in a table is not important and should have no effect on the results of any query.
- The order of the fields relative to each other is not important. Conventional practice is that the unique identifier field that identifies each record (the field usually called *ID*) is the first field in the table.
- Each record should be unique in that it differs from all the other records in the table. The record may differ on only one field such as the *ID* field, but nonetheless each row (record) should be unique. In a table of library books, a library with two identical books should be able to distinguish between the two books by a field called

8 ■ Using Access in Forensic Investigations

Copy (or something similar) and the first copy of the book could have *Copy* = 1 and the second copy of the book could have *Copy* = 2.

- A table should have a primary key that is unique and that contains no duplicate values so that each record (row) can be identified uniquely. A table can also have a foreign key, which is a way to link to the primary key in another table.
- The field values must pertain to the subject matter of the table and must completely describe the contents of the table. A table for library books should hold all the data pertaining to each book, and should not contain superfluous data such as the home address of the last patron to read the book.
- The preferred situation is that users should be able to change the data in one field without affecting any of the other fields. Access 2010 does allow users to have a calculated data type. This means that, for example, *ExtendedValue* could be equal to *Count* * *Amount*. If either *Count* or *Amount* is updated, then *ExtendedValue* is updated automatically.

If the data for the investigation is already in an Access format then the analysis can begin with little or no data preparation. When the data is in the form of a flat file (or files) then the data needs to be imported into Access. Some preparation work is also needed when the database was created in a prior version of Access. These prior-version databases can be converted to Access 2007 databases. The new Access 2007 file format is preferred because it has some new functions that were not previously available. Access 2007 is backward-compatible to Access 97.

■ IMPORTING DATA INTO ACCESS

Importing data into Access is reasonably straightforward. Data is imported from Excel using **External Data**→**Import**→**Excel** as is shown in Figure 1.6.

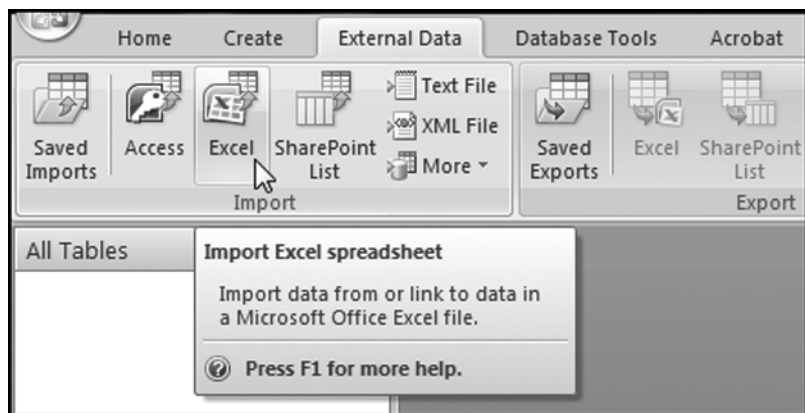


FIGURE 1.6 Commands Used to Import Data from Excel into Access

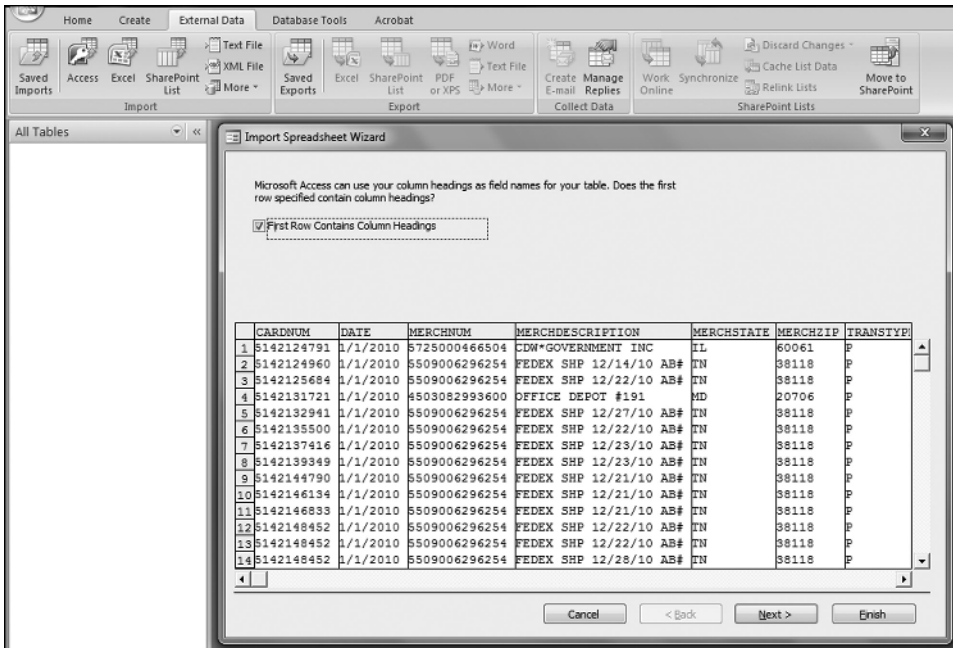


FIGURE 1.7 Import Spreadsheet Wizard Used to Import Data from Excel

Figure 1.6 shows the starting steps for importing data from Excel. Exporting data and results from Access to Excel can present some challenges when the data exceeds the size of the clipboard. One solution is to then use Excel to import the data from Access. The Import Spreadsheet Wizard for importing data from Excel is shown in Figure 1.7.

Importing data one sheet at a time from Excel is reasonably straightforward. It makes the importing procedure easier if the first row in Excel contains column headings. It is usually a good idea to format any field that will be used for calculations as the *Currency* data type. The imported data is shown in Figure 1.8.

Purchasing card data is shown in Figure 1.8 in a table that looks like a familiar Excel worksheet. A difference between Access and Excel is that in Access all calculations need

ID	CARDNUM	DATE	MERCHNUM	MERCHDESC	MERCHSTAT	MERCHZIP	TRANSTYP	AMOUNT
1	5142124791	1/1/2010	5725000466504	CDW*GOVERN	IL	60061	P	106.89
2	5142124960	1/1/2010	5509006296254	FEDEX SHP 12/	TN	38118	P	3.62
3	5142125684	1/1/2010	5509006296254	FEDEX SHP 12/;	TN	38118	P	3.67
4	5142131721	1/1/2010	4503082993600	OFFICE DEPOT	MD	20706	P	178.49
5	5142132941	1/1/2010	5509006296254	FEDEX SHP 12/;	TN	38118	P	3.62
6	5142135500	1/1/2010	5509006296254	FEDEX SHP 12/;	TN	38118	P	3.62
7	5142137416	1/1/2010	5509006296254	FEDEX SHP 12/;	TN	38118	P	3.74
8	5142139349	1/1/2010	5509006296254	FEDEX SHP 12/;	TN	38118	P	3.67
9	5142144790	1/1/2010	5509006296254	FEDEX SHP 12/;	TN	38118	P	11.29
10	5142146134	1/1/2010	5509006296254	FEDEX SHP 12/;	TN	38118	P	3.85
11	5142146833	1/1/2010	5509006296254	FEDEX SHP 12/;	TN	38118	P	3.62
12	5142148452	1/1/2010	5509006296254	FEDEX SHP 12/;	TN	38118	P	3.74
13	5142148452	1/1/2010	5509006296254	FEDEX SHP 12/;	TN	38118	P	3.74
14	5142148452	1/1/2010	5509006296254	FEDEX SHP 12/;	TN	38118	P	3.62
15	5142148452	1/1/2010	5509006296254	FEDEX SHP 12/;	TN	38118	P	3.62

FIGURE 1.8 Purchasing Card Data in Excel

10 ■ Using Access in Forensic Investigations

to be done using queries. Another difference is that (almost) all changes to tables such as edits to records, deletions of records, additions of records, and deletions of fields are permanent. Excel has the **Control+Z** command to backtrack, but in Access there is no option to either backtrack or to exit without saving.

A REVIEW OF ACCESS QUERIES

Queries are the main focus in forensic analytics. A query is essentially a question, and forensic analytics is all about asking questions and scrutinizing or auditing the answers. The main types of queries are reviewed below:

- **Creating calculated fields.** Here we create one or more fields in the table that are calculated values using the data in the other fields. For example, with Benford's Law we need to calculate the first-two digits in every number and this first step is a query. The general rule is that any calculation is always based on other field values in that same record. For example, quantity times unit price will give us a total cost. Access can easily perform calculations using field values from the same row or record. It is difficult to perform a calculation that requires Access to use a field value from a preceding or succeeding row. An example of such a calculation is a cumulative sum. The problem with using preceding or succeeding rows is that if the table is resorted then the cumulative sums need to be recalculated and the order of the records in a table should not affect a calculated value.
- **Grouping records.** In these queries various parameters are calculated for each group in a field (e.g., *CardNum*, *MerchNum*, *Date*, or *MerchZip*). Examples of these parameters are the sum, average, count, maximum, minimum, first, last, or the standard deviation. Some forensic analytics tests simply involve calculating the sums or averages for selected groups of records.
- **Identifying duplicate records.** In these queries duplicate records are identified. This will usually be a selective identification of duplicates because one of the criteria in table design is that all the records are unique. This query will usually look for cases where we have duplicates on two or three fields only.
- **Filtering data.** Access has a powerful filtering function and many types of conditions can be used. A query could be used to show all the purchasing card transactions for employee *x* for a range of dates (perhaps a range when the employee was on vacation). The filter could be combined with a grouping command using the powerful *Where* criteria in Access.
- **Using a Join to query conditions in two or more tables.** A query that requires Access to use the data in two or more tables needs to include a *Join*. The most common type of Join is where we identify all our forensic units of interest at the start of the analysis and we want the next query to only give us the results for our selected vendors, merchants, or employees.
- **Appending data.** Append queries are important in forensic analytics because these queries can be used to retrieve data from one table and add it to another table.

This is a useful way to add (say) November's data to the year-to-date data table. Append queries are also useful to convert data from an Excel format where the data for each time period is in separate columns, to the table format in an Access database where the data for the various time periods are stacked on each other. An example is shown later in this chapter.

- **Crosstab queries.** Crosstab queries allow users to add another level of grouping. With the purchasing card data one could calculate the merchant totals for the year. A crosstab query could also add another layer of analysis to also include merchant totals per month.
- **Parameter query.** A parameter query returns all the records for a specified field value. This is useful for the risk-scoring models in Chapters 15, 16, and 17. A parameter query would be used to show all the card transactions for the *Crown Plaza Hotel* as is shown in Figure 1.9.

Figure 1.9 shows a parameter query in Design View. The “Enter Name of Merchant” in square brackets is an informative message that appears when the query is run. The query is run by clicking **Design**→**Results**→**Run**, and the dialog screen is shown in Figure 1.10.

Figure 1.10 shows the dialog box of a parameter query. The words *Crown Plaza Hotel* are entered and after clicking **OK** the results will show only the transactions for the Crown Plaza Hotel. A parameter query can have more than one parameter.

Queries are the workhorses of forensic analytics and the book shows many examples of queries from Chapter 4 through Chapter 18. Reports are either based on tables or queries. In a forensic environment the reports will usually be based on queries. The only real issue with Access is with calculations that are based on records that come before or after the record in question. Access has difficulty in looking up and down when performing calculations.

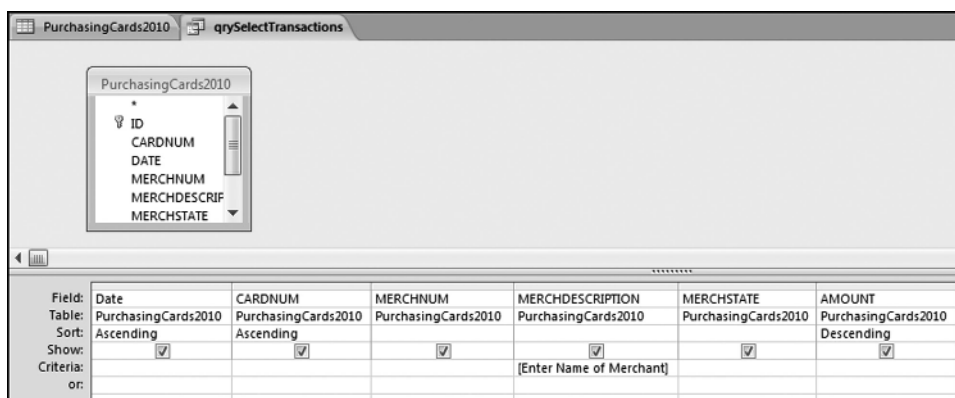


FIGURE 1.9 Parameter Query in Design View. The Query Is a Parameter Query Because of the “Enter Name of Merchant” in Square Brackets

12 ■ Using Access in Forensic Investigations

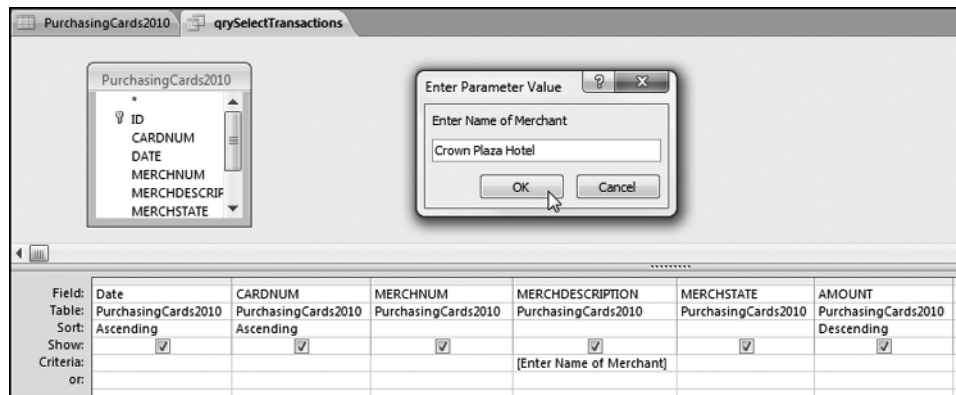


FIGURE 1.10 Dialog Box of a Parameter Query

Some forensic analytics tests will use several queries to get the final result. The general format for a query is to state which table (or tables or prior query) should be used, which fields (or columns) are to be queried, what calculations or comparisons should be done, which records should be returned, and how to format the output (sorting is one option). The usual method will be to use the **Create**→**Other**→**Query Design** sequence to create queries. The important features in Access supporting queries are:

- The ability to create queries using the wizards, Design View, or SQL view.
- The ability to query a combination of one or more tables or the results of prior queries.
- The ability to use SQL to change a query created in Design View.
- The Performance Analyzer (**Database Tools**→**Analyze**→**Analyze Performance**), which helps to make queries more efficient.
- The ability to format the output of the query (usually by displaying results to two digits after the decimal point).
- The ability to sort and resort query results without creating more queries.
- The extensive library of built-in functions for calculated fields.
- The built-in statistical operations such as Sum, Count, Average, Minimum, Maximum, First, and Last.
- The built-in If (Immediate If) function and the Switch function, which allows for multiple If statements, together with a full complement of operators including **And**, **Or**, and **Not**.
- The ability to work with empty (null) fields.
- The ability to easily export tables and the results of queries to Excel for further analysis or neat presentation.

Access was made to analyze data and the calculation speed is quite remarkable. With practice and patience the Access grid becomes quite logical. The next section demonstrates how to prepare Excel data for use in Access.

U.S. No. 2 Fuel Oil All Sales/Deliveries by Prime Supplier (Thousand Gallons per Day)												
Year	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
1983	65,251.4	63,848.0	62,375.4	53,049.0	36,302.8	40,598.8	35,802.6	41,119.8	38,150.2	41,636.8	54,958.2	78,119.2
1984	88,893.7	66,567.9	74,816.9	54,803.4	48,251.7	40,018.0	40,495.9	45,299.1	47,100.2	46,981.7	61,917.7	74,286.6
1985	87,392.1	77,660.0	67,731.4	55,829.4	45,673.8	40,889.3	41,082.0	44,985.2	45,262.7	54,954.8	58,369.5	80,912.0
1986	72,150.2	80,349.8	64,523.4	51,892.4	47,643.8	40,559.1	40,221.2	47,773.9	46,787.6	55,272.2	58,240.2	74,448.9
1987	78,148.0	73,292.2	65,579.3	52,600.1	46,028.2	43,426.8	44,110.6	41,079.7	47,423.7	60,538.5	62,190.7	78,371.5
1988	90,618.3	89,824.4	77,098.3	56,688.9	50,735.3	47,457.9	48,687.8	50,874.4	50,453.2	61,333.0	67,826.8	82,290.2
1989	79,098.6	79,580.9	75,144.4	55,264.5	46,260.3	45,547.3	39,889.0	50,778.4	50,742.5	56,433.0	64,378.7	92,250.9
1990	74,256.8	73,752.3	63,254.8	55,571.5	49,431.4	49,050.2	46,314.1	54,196.0	53,803.8	54,397.2	60,710.9	64,933.7
1991	81,366.5	68,472.4	60,802.5	53,574.6	46,538.5	42,699.6	43,889.8	46,724.4	46,882.3	53,013.5	55,807.4	68,245.1
1992	78,609.0	70,094.2	64,700.2	57,430.5	46,793.0	46,859.4	43,287.0	44,860.2	52,209.5	56,162.9	59,294.2	72,803.0
1993	62,959.1	69,927.9	65,619.9	49,032.0	38,744.9	41,405.4	38,368.4	40,131.1	41,414.2	32,073.2	38,875.9	51,627.7
1994	68,199.7	62,191.5	49,804.4	32,960.3	27,444.2	26,009.7	22,623.7	27,619.0	30,391.2	32,976.0	34,185.7	45,932.5
1995	47,443.4	54,852.4	41,434.2	31,347.5	26,195.3	24,322.2	21,676.0	24,859.5	26,625.0	29,076.7	38,986.2	53,054.9
1996	57,988.5	55,685.0	42,789.1	33,643.2	25,651.8	23,238.7	22,987.0	24,188.8	27,810.4	31,993.5	38,958.1	44,723.2
1997	52,887.5	46,281.0	39,457.7	33,128.7	25,986.1	24,433.3	23,436.9	23,967.5	27,152.4	30,358.4	35,005.9	45,200.0
1998	43,253.5	42,453.9	38,313.8	28,122.4	22,565.9	23,580.9	22,198.4	21,111.8	22,821.0	26,745.2	30,257.7	35,790.4
1999	45,674.2	42,867.5	40,885.0	25,178.0	20,491.4	19,771.7	19,444.3	20,478.6	20,686.4	26,551.8	31,225.7	39,182.7
2000	44,302.7	44,854.6	32,556.5	26,536.6	23,096.5	21,336.7	18,980.8	23,288.6	25,079.8	28,444.4	32,517.4	45,990.5
2001	53,025.9	47,172.5	41,607.6	31,383.8	22,890.0	20,873.6	20,666.7	21,538.4	22,254.6	26,139.6	29,756.8	34,409.8
2002	41,101.6	37,246.4	31,883.4	25,488.2	21,412.1	18,680.5	19,014.4	19,157.7	20,522.8	26,888.9	32,496.5	42,760.4
2003	48,363.6	44,799.6	35,392.1	27,697.4	20,906.8	18,965.7	18,912.4	19,226.8	22,797.1	26,763.4	26,669.6	37,965.0
2004	47,979.8	41,852.4	32,449.3	25,977.2	18,251.6	17,929.4	16,702.6	18,643.1	20,138.4	23,311.7	27,318.4	35,468.5
2005	38,673.1	37,556.4	33,096.5	21,996.3	19,922.3	18,032.3	15,622.7	18,866.3	19,670.4	20,815.7	25,978.4	34,780.1
2006	31,740.5	33,940.7	30,423.9	18,602.6	15,026.0	13,199.1	12,192.2	13,805.4	14,157.7	16,865.6	20,107.2	22,413.7
2007	28,370.3	35,872.9	26,326.4	18,745.5	12,724.6	9,494.8	8,344.7	8,942.9	8,861.3	10,809.0	17,532.6	24,448.1
2008	26,658.0	26,270.8	20,195.3	13,160.3	8,227.9	7,024.7	6,553.1	7,541.1	9,002.4	12,898.3	15,728.3	24,143.0
2009	29,853.1	25,136.2	19,550.3	12,196.2	7,314.3	6,297.5	5,262.4	5,057.2	7,518.7	10,975.7	11,797.9	20,549.5
2010	24,449.1	23,328.9	13,491.6	6,992.0	5,325.0	4,545.6	3,314.7	4,381.5	5,529.0	8,588.5		

FIGURE 1.11 U.S. Fuel Oil Sales from 1983 to 2010

CONVERTING EXCEL DATA INTO A USABLE ACCESS FORMAT

Data tables that are developed in Excel usually do not follow the rules and logic of database tables. These Excel tables need to be “converted” to a usable Access format. Quite often these Access conversions need to be performed on data downloaded from statistical agencies. An example of such a table is the Fuel Oil table of the EIA shown in Figure 1.11. This data was copied from the U.S. Energy Information Administration’s website (www.eia.gov) by clicking through to **Petroleum**→**Prime Supplier Sales Volumes**→**No. 2 Fuel Oil** (1983–2010).

The fuel oil data in Figure 1.11 is accumulated row by row. As time progresses, more rows are added to the bottom of the table. In other Excel worksheets columns could be added to the right of the table as time progresses. This data was imported into Excel using the **Copy** and **Paste** commands. A portion of the Excel file is shown in Figure 1.12.

This data needs some preparatory steps because Access cannot work with time-related data when the time period is indicated in the field’s name (e.g., Jan, Feb, or Mar).

14 ■ Using Access in Forensic Investigations

Year	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
1983	65,251.40	63,848.00	62,375.40	53,049.00	36,302.80	40,598.80	35,802.60	41,119.80	38,150.20	41,636.80	54,958.20	78,119.20
1984	88,893.70	66,567.90	74,816.90	54,803.40	48,251.70	40,018.00	40,495.90	45,299.10	47,100.20	46,981.70	61,917.70	74,286.60
1985	87,392.10	77,660.00	67,731.40	55,829.40	45,673.80	40,889.30	41,082.00	44,985.20	45,262.70	54,954.80	58,369.50	80,912.00
1986	72,150.20	80,349.80	64,523.40	51,892.40	47,643.80	40,559.10	40,221.20	47,773.90	46,787.60	55,272.20	58,240.20	74,448.90
1987	78,148.00	73,292.20	65,579.30	52,600.10	46,028.20	43,426.80	44,110.60	41,079.70	47,423.70	60,538.50	62,190.70	78,371.50
1988	90,618.30	89,824.40	77,098.30	56,688.90	50,735.30	47,457.90	48,687.80	50,874.40	50,453.20	61,333.00	67,826.80	82,290.20
1989	79,098.60	79,580.90	75,144.40	55,264.50	46,260.30	45,547.30	39,889.00	50,778.40	50,742.50	56,433.00	64,378.70	92,250.90
1990	74,256.80	73,752.30	63,254.80	55,571.50	49,431.40	49,050.20	46,314.10	54,196.00	53,803.80	54,397.20	60,710.90	64,933.70
1991	81,366.50	68,472.40	60,802.50	53,574.60	46,538.50	42,699.60	43,889.80	46,724.40	46,882.30	53,013.50	55,807.40	68,245.10
1992	78,609.00	70,094.20	64,700.20	57,430.50	46,793.00	46,859.40	43,287.00	44,860.20	52,209.50	56,162.90	59,294.20	72,803.00
1993	62,959.10	69,927.90	65,619.90	49,032.00	38,744.90	41,405.40	38,368.40	40,131.10	41,414.20	32,073.20	38,875.90	51,627.70
1994	68,199.70	62,191.50	49,804.40	32,960.30	27,444.20	26,009.70	22,623.70	27,619.00	30,391.20	32,976.00	34,185.70	45,932.50
1995	47,443.40	54,852.40	41,434.20	31,347.50	26,195.30	24,322.20	21,676.00	24,859.50	26,625.00	29,076.70	38,986.20	53,054.90
1996	57,988.50	55,685.00	42,789.10	33,643.20	25,651.80	23,238.70	22,987.00	24,188.80	27,810.40	31,993.50	38,958.10	44,723.20

FIGURE 1.12 Fuel Oil Data in an Excel Worksheet

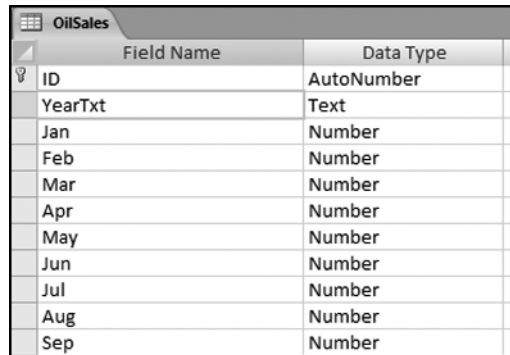
Many types of Excel layouts exist and they all need to be converted to an Access-friendly format. The blank rows can be deleted by highlighting the blank rows one at a time and then deleting the row because we only have six blank rows. Another option would be to sort the Excel table so that all the blanks are at the top of the table and then to delete the blank rows. You might need to copy the smaller table to a new Excel worksheet before importing this into Excel. This is because Excel seems to remember that the original table had (say) 35 rows and when it is imported into Access then Access imports 35 rows, even though the last six rows are blank. The Access table is shown in Figure 1.13.

Figure 1.13 shows the Access table with the Excel fuel oil data. The first step is to use Design View to change the name of the field *Year* to *YearTxt* (for year text). This is because the new table will have a field called *Year* with *Year* being a numeric field. The name change is shown in Figure 1.14.

The field name is changed to *YearTxt* in Design View in Figure 1.14. The table can now be converted to an Access format. The next step is to convert the numeric values to

ID	Year	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
1	1983	65,251.40	63,848.00	62,375.40	53,049.00	36,302.80	40,598.80	35,802.60	41,119.80	38,150.20	41,636.80	54,958.20	78,119.20
2	1984	88,893.70	66,567.90	74,816.90	54,803.40	48,251.70	40,018.00	40,495.90	45,299.10	47,100.20	46,981.70	61,917.70	74,286.60
3	1985	87,392.10	77,660.00	67,731.40	55,829.40	45,673.80	40,889.30	41,082.00	44,985.20	45,262.70	54,954.80	58,369.50	80,912.00
4	1986	72,150.20	80,349.80	64,523.40	51,892.40	47,643.80	40,559.10	40,221.20	47,773.90	46,787.60	55,272.20	58,240.20	74,448.90
5	1987	78,148.00	73,292.20	65,579.30	52,600.10	46,028.20	43,426.80	44,110.60	41,079.70	47,423.70	60,538.50	62,190.70	78,371.50
6	1988	90,618.30	89,824.40	77,098.30	56,688.90	50,735.30	47,457.90	48,687.80	50,874.40	50,453.20	61,333.00	67,826.80	82,290.20
7	1989	79,098.60	79,580.90	75,144.40	55,264.50	46,260.30	45,547.30	39,889.00	50,778.40	50,742.50	56,433.00	64,378.70	92,250.90
8	1990	74,256.80	73,752.30	63,254.80	55,571.50	49,431.40	49,050.20	46,314.10	54,196.00	53,803.80	54,397.20	60,710.90	64,933.70
9	1991	81,366.50	68,472.40	60,802.50	53,574.60	46,538.50	42,699.60	43,889.80	46,724.40	46,882.30	53,013.50	55,807.40	68,245.10
10	1992	78,609.00	70,094.20	64,700.20	57,430.50	46,793.00	46,859.40	43,287.00	44,860.20	52,209.50	56,162.90	59,294.20	72,803.00
11	1993	62,959.10	69,927.90	65,619.90	49,032.00	38,744.90	41,405.40	38,368.40	40,131.10	41,414.20	32,073.20	38,875.90	51,627.70
12	1994	68,199.70	62,191.50	49,804.40	32,960.30	27,444.20	26,009.70	22,623.70	27,619.00	30,391.20	32,976.00	34,185.70	45,932.50
13	1995	47,443.40	54,852.40	41,434.20	31,347.50	26,195.30	24,322.20	21,676.00	24,859.50	26,625.00	29,076.70	38,986.20	53,054.90
14	1996	57,988.50	55,685.00	42,789.10	33,643.20	25,651.80	23,238.70	22,987.00	24,188.80	27,810.40	31,993.50	38,958.10	44,723.20
15	1997	52,887.50	46,281.00	39,457.70	33,128.70	25,986.10	24,433.30	23,436.90	23,967.50	27,152.40	30,358.40	35,005.90	45,200.00

FIGURE 1.13 The Access Table with the Imported Excel Fuel Oil Data

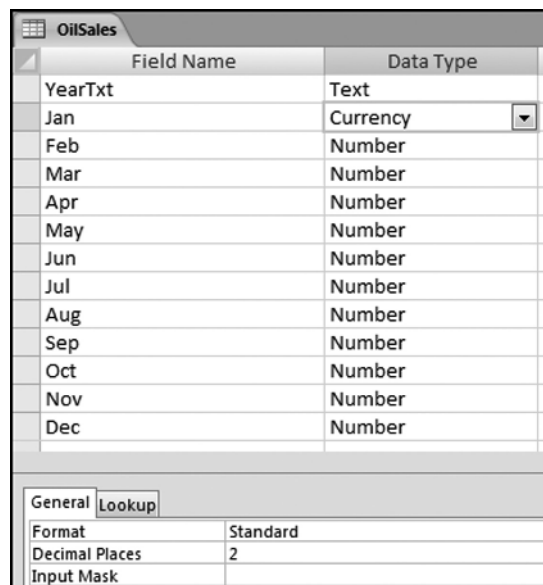


Field Name	Data Type
ID	AutoNumber
YearTxt	Text
Jan	Number
Feb	Number
Mar	Number
Apr	Number
May	Number
Jun	Number
Jul	Number
Aug	Number
Sep	Number

FIGURE 1.14 Field Name Changed to Yeartxt in Design View

Currency. It is best to do this conversion at this early stage. The Currency conversions need to be done for each of the 12 numeric fields and the first conversion is shown in Figure 1.15.

This conversion needs to be done for all 12 numeric fields. The table needs to be saved before the changes take effect. Access gives a prompt that some accuracy might be lost with the currency format. When the table is viewed again in Datasheet View, the numbers will usually (but not always) be shown with leading dollar signs and negative numbers in parentheses. The currency format helps to prevent rounding errors in calculations.



Field Name	Data Type
YearTxt	Text
Jan	Currency
Feb	Number
Mar	Number
Apr	Number
May	Number
Jun	Number
Jul	Number
Aug	Number
Sep	Number
Oct	Number
Nov	Number
Dec	Number

General	Lookup
Format	Standard
Decimal Places	2
Input Mask	

FIGURE 1.15 Conversion of the Field Jan to Currency with Two Decimal Places

16 ■ Using Access in Forensic Investigations



FIGURE 1.16 Make Table Query Used to Start the Process of Building a New Access Table

The next step is to create a table that will be the starting building block for our complete table. This is done with a Make Table query as is shown in Figure 1.16. The January data is used as a foundation to start the ball rolling. The new table is called *OilSales2*.

The conversion of a text field to a numeric value is sometimes tricky. In this case the *Year* field had two spaces to the left of the visible characters, which is not usually an issue with data formatted as text. The conversion to a numeric value required the use of the *Val* (value) and the *Mid* (middle) functions as shown below:

$$\text{Year} : \text{Val}(\text{Mid}([\text{YearTxt}], 3, 4))$$

The field *Month* was converted from *Jan* to the number *1*, which makes it easier to use in queries. The *GallonsPD* (gallons per day) field was formatted as currency using the field properties. The *GallonsPM* (gallons per month) field was automatically formatted as currency. The table is in gallons per day and the new table will include both the daily average and the monthly total. Even though **OK** is clicked in the dialog box in Figure 1.16, the query must still be run using **Design**→**Results**→**Run**. Access always gives a warning that you are about to paste *x* rows into a new table. This warning can be ignored if you are safely below the size limit of an Access database. Click **Yes** and the *OilSales2* table should be as is shown in Figure 1.17.

The next step is to Append the February data to this table and then to do the same for all the other months. The query to append February is shown in Figure 1.18.

The fields and data from *OilSales* are appended to *OilSales2*. The monthly total is a little complex because February sometimes has 28 days and sometimes the month has 29 days. The formula for *GallonsPM* is:

$$\begin{aligned} \text{GallonsPM} &: \text{Iif}([\text{Year}] = 1984 \text{ Or } [\text{Year}] = 1988 \text{ Or } [\text{Year}] \\ &= 1992 \text{ Or } [\text{Year}] = 1996 \text{ Or } [\text{Year}] = 2000 \text{ Or } [\text{Year}] \\ &= 2004 \text{ Or } [\text{Year}] = 2008, [\text{Feb}] * 29, [\text{Feb}] * 28) \end{aligned}$$

Year	Month	GallonsPD	GallonsPM
1983	1	\$65,251.40	\$2,022,793.40
1984	1	\$88,893.70	\$2,755,704.70
1985	1	\$87,392.10	\$2,709,155.10
1986	1	\$72,150.20	\$2,236,656.20
1987	1	\$78,148.00	\$2,422,588.00
1988	1	\$90,618.30	\$2,809,167.30
1989	1	\$79,098.60	\$2,452,056.60
1990	1	\$74,256.80	\$2,301,960.80
1991	1	\$81,366.50	\$2,522,361.50
1992	1	\$78,609.00	\$2,436,879.00
1993	1	\$62,959.10	\$1,951,732.10
1994	1	\$68,199.70	\$2,114,190.70
1995	1	\$47,443.40	\$1,470,745.40
1996	1	\$57,988.50	\$1,797,643.50
1997	1	\$52,887.50	\$1,639,512.50
1998	1	\$43,253.50	\$1,340,858.50
1999	1	\$45,674.20	\$1,415,900.20
2000	1	\$44,302.70	\$1,373,383.70
2001	1	\$53,025.90	\$1,643,802.90
2002	1	\$41,101.60	\$1,274,149.60
2003	1	\$48,363.60	\$1,499,271.60
2004	1	\$47,979.80	\$1,487,373.80
2005	1	\$38,673.10	\$1,198,866.10
2006	1	\$31,740.50	\$983,955.50
2007	1	\$28,370.30	\$879,479.30

FIGURE 1.17 The First Table in the Creation of *OilSales2*

The screenshot shows the 'Append' query design grid in Microsoft Access. The design grid is as follows:

Field:	Year: Val(Mid([YearTxt],3,4))	Month: 2	GallonsPD: Feb OilSales	GallonsPM: IIf([Year]=1984 Or [Year]=1988 Or [Year]=1992 Or [Year]=
Table:				
Sort:				
Append To:	Year	Month	GallonsPD	GallonsPM
Criteria:				

FIGURE 1.18 The Append Query Used to Build the *OilSales2* Table

18 ■ Using Access in Forensic Investigations

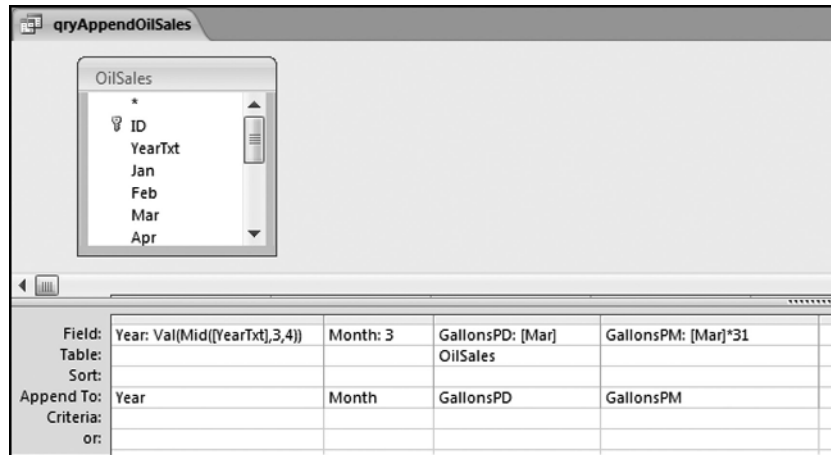


FIGURE 1.19 Query Used to Append the March Data

The formula uses the If function (IIf in Access for Immediate if) together with the Or function.

The query needs to be run using **Design**→**Results**→**Run**. Access gives a warning that you are about to append 28 rows. Once you have clicked **Yes**, the command cannot be undone. Run the query and click **Yes**. It is a good idea to make backup copies of your tables until you are quite familiar with the appending process. The query used for appending the March data is shown in Figure 1.19.

The Month is changed to “3” without any quotes, and the gallons per day and gallons per month formulas are also revised. The *GallonsPM* calculation for March is simply the gallons per day multiplied by 31. There is no leap year complication. This process is repeated for March through December. The final table is shown in Figure 1.20.

The record indicator at the bottom of the screen shows that there are 336 records in the table. This is correct because there are 28 years and 28*12 months equals 336 records. Access does not necessarily stack the tables one on top of the other in the order in which the append queries were run. One way to tidy up the table is to use another Make Table query to sort the data as you would like it to be sorted. It is good practice to check whether each month has been added just once. One or two queries can confirm this and the query in Figure 1.21 counts and sums the records for each month.

The query in Figure 1.21 tests whether there are 27 or 28 records per year and also whether the average of the numbers is logical. The results are shown in Figure 1.22.

The results of the query in Figure 1.22 confirm that the appending steps were done correctly. For each month there are either 27 or 28 records. September to December, 2010, did not have data at the time that the file was downloaded and the results show that months 9 to 12 have only 27 records. The average gallons per day has a seasonal pattern with high sales in the cold winter months (12, 1, 2, and 3 corresponding to December to March) and low sales in the summer months (5 to 8 corresponding to May to August). The table *OilSales2* can now be used for Access queries. This heating oil example is continued in Chapter 14 with the heating oil sales application.

Year	Month	GallonsPD	GallonsPM
1983	1	\$65,251.40	\$2,022,793.40
1984	1	\$88,893.70	\$2,755,704.70
1985	1	\$87,392.10	\$2,709,155.10
1986	1	\$72,150.20	\$2,236,656.20
1987	1	\$78,148.00	\$2,422,588.00
1988	1	\$90,618.30	\$2,809,167.30
1989	1	\$79,098.60	\$2,452,056.60
1990	1	\$74,256.80	\$2,301,960.80
1991	1	\$81,366.50	\$2,522,361.50
1992	1	\$78,609.00	\$2,436,879.00
1993	1	\$62,959.10	\$1,951,732.10
1994	1	\$68,199.70	\$2,114,190.70
1995	1	\$47,443.40	\$1,470,745.40
1996	1	\$57,988.50	\$1,797,643.50
1997	1	\$52,887.50	\$1,639,512.50
1998	1	\$43,253.50	\$1,340,858.50
1999	1	\$45,674.20	\$1,415,900.20
2000	1	\$44,302.70	\$1,373,383.70
2001	1	\$53,025.90	\$1,643,802.90
2002	1	\$41,101.60	\$1,274,149.60
2003	1	\$48,363.60	\$1,499,271.60
2004	1	\$47,979.80	\$1,487,373.80
2005	1	\$38,673.10	\$1,198,866.10
2006	1	\$31,740.50	\$983,955.50
2007	1	\$28,370.30	\$879,479.30

Record: 1 of 336 No Filter Search

FIGURE 1.20 Completed Heating Oil Table

qryCheckResults			
OilSales2			
*			
Year			
Month			
GallonsPD			
GallonsPM			
Field:	Month	GallonsPD	GallonsPD
Table:	OilSales2	OilSales2	OilSales2
Total:	Group By	Count	Avg
Sort:			
Show:	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
Criteria:			

FIGURE 1.21 How to Check Whether the Append Queries Were Correctly Run

20 ■ Using Access in Forensic Investigations

Month	CountOfGallonsPD	AvgOfGallonsPD
1	28	\$56,739.58
2	28	\$54,133.31
3	28	\$46,832.27
4	28	\$36,031.88
5	28	\$29,494.13
6	28	\$27,723.15
7	28	\$26,431.48
8	28	\$28,942.81
9	27	\$31,322.94
10	27	\$35,348.54
11	27	\$40,410.84
12	27	\$52,040.78

FIGURE 1.22 Results of the Query Designed to Test the Appending Operations

USING THE ACCESS DOCUMENTER

A forensic report is prepared after a forensic investigation is completed. This report should describe all the evidence gathered, the findings, conclusions, recommendations, and the corrective actions (if any) that were taken. The contents of this report should have a tone that is not inflammatory, libelous, or with prejudicial connotations. The report should include a description of the forensic analytics work that was done. The working papers should include a copy of the data analyzed on either a CD or a USB flash drive, and the results of the queries. A full description of the database should also be included in the report. A useful feature in Access is the **Database Documenter**. The database documenter is activated by using **Database Tools**→**Analyze**→**Database Documenter**. The dialog screen is shown in Figure 1.23.

For a complete documentation each object (in this case just Tables and Queries) needs to be selected using **Select All**. Click **OK** to run the documenter. The documentation is comprehensive and includes facts related to the database objects and the SQL code describing the queries. With the SQL code, the same query can be run on another computer using the same data table. The documenter also includes the time and date that the table was last updated giving a record of any changes to the table after a query was run. The Database Documenter does not meet the standards of absolute proof but it goes a long way to documenting and supporting a description of the tests that were run.

Another useful Access feature is the ability to describe tables and queries in the table and query properties. The **Table Properties** dialog box is activated by right clicking on the table names to give the dialog box shown in Figure 1.24.

The table description is entered using the Table Properties dialog box shown in Figure 1.24. The **Apply** and **OK** buttons are used after the description is typed. The fields can be described when the table is in Design View as is the case in Figures 1.14

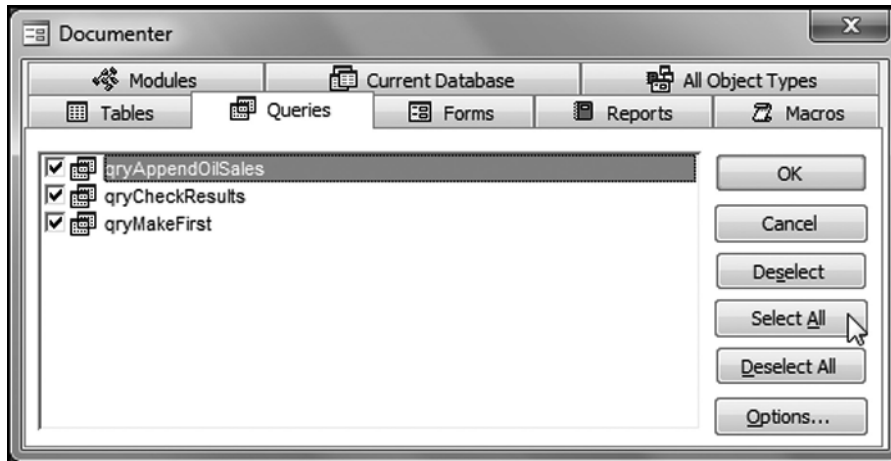


FIGURE 1.23 Dialog Screen for the Database Documenter

and 1.15. Access also allows for a complete description to be included for all queries. The **Query Properties** dialog box is activated using a right click on the query name and clicking **Object Properties**. An example is shown in Figure 1.25.

Access allows for a reasonably long description of each query using the Object Properties shown in Figure 1.25. The buttons **Apply** and **OK** are used to save the description. There is also a way to include a detailed description of the whole database using **Manage**→**Database Properties** as is shown in Figure 1.26.

The step to retrieve the database properties is shown in Figure 1.26. The details are shown across five tabs. A printout or an electronic jpg image of each of the tabs should be included in the working papers. The **Contents** tab is shown in Figure 1.27.

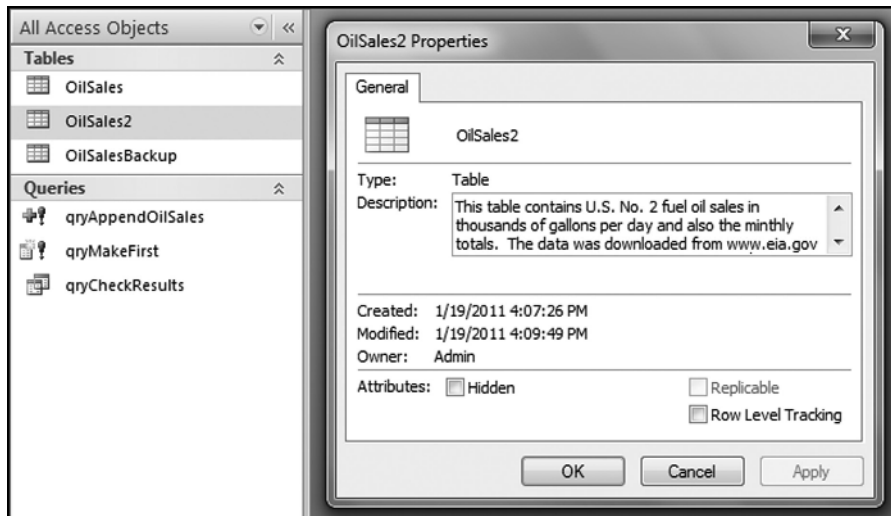


FIGURE 1.24 Dialog Box Used to Enter the Table Description

22 ■ Using Access in Forensic Investigations

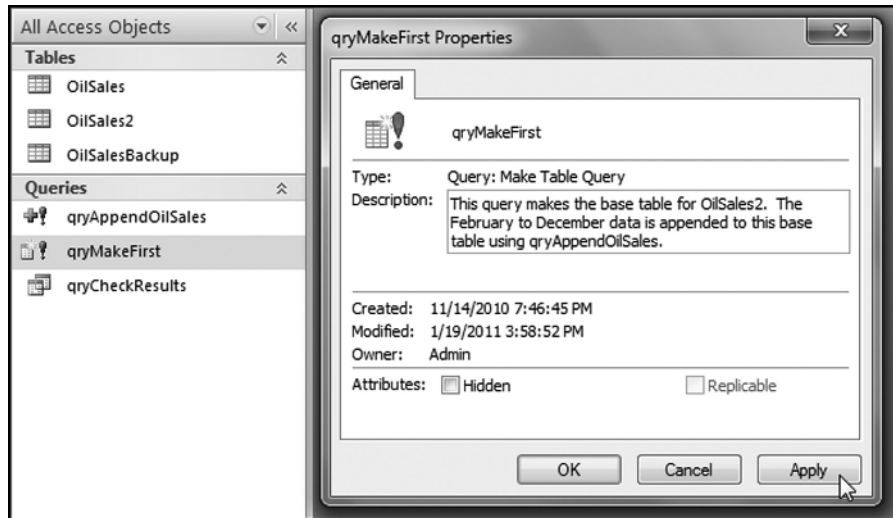


FIGURE 1.25 Dialog Box Used to Include a Description of a Query

The **Contents** tab lists the names of all the Access objects. The **Summary** tab is made up of details added by the forensic analyst. The database properties together with the documenter, the descriptions that can be included in the Design View of a table, and the tables and queries properties all make it easier for the analyst, or someone else, to understand the contents of the database. The table and query properties can be seen by expanding the details shown in the Navigation Pane. The procedure to see the properties is shown in Figure 1.28.

The procedure to view the object details is to right click on either the Tables or the Queries heading and then select **View By**→**Details**. The details will then be visible in the Navigation Pane. To return to the names only one would select **List**. The documentation options are valuable and allow other users to understand the contents when the database is used at some time in the future.

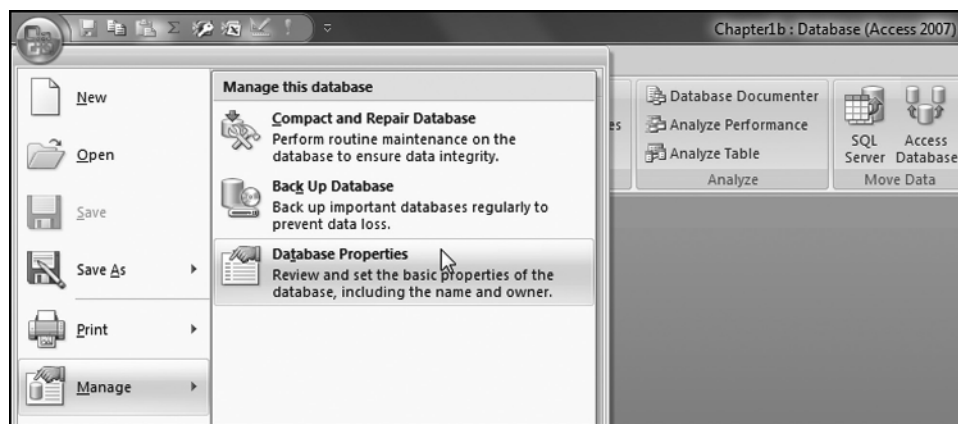


FIGURE 1.26 Retrieving the *Database Properties* Options in Access

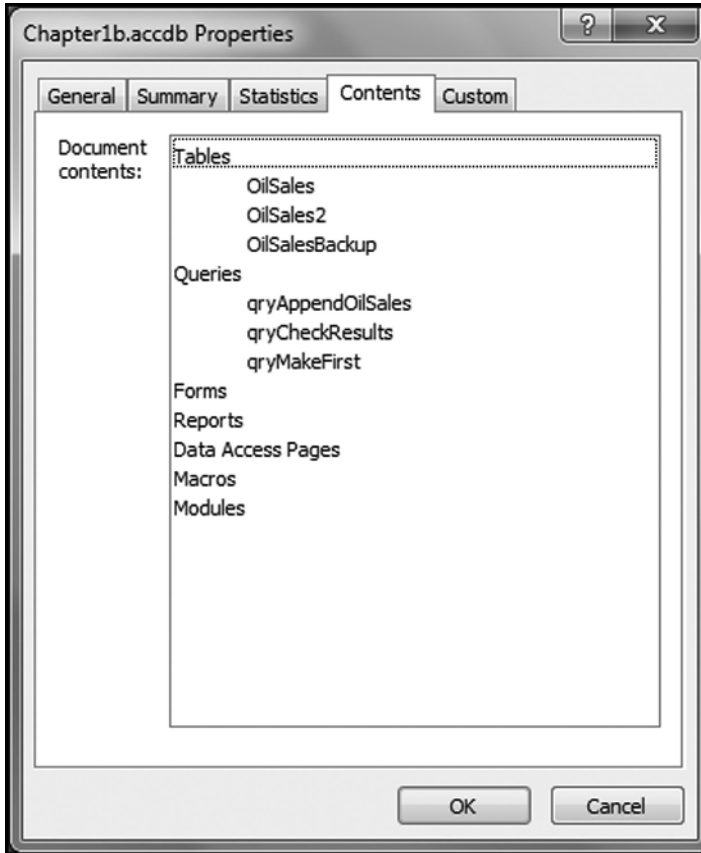


FIGURE 1.27 Database Properties Documentation Feature of Access

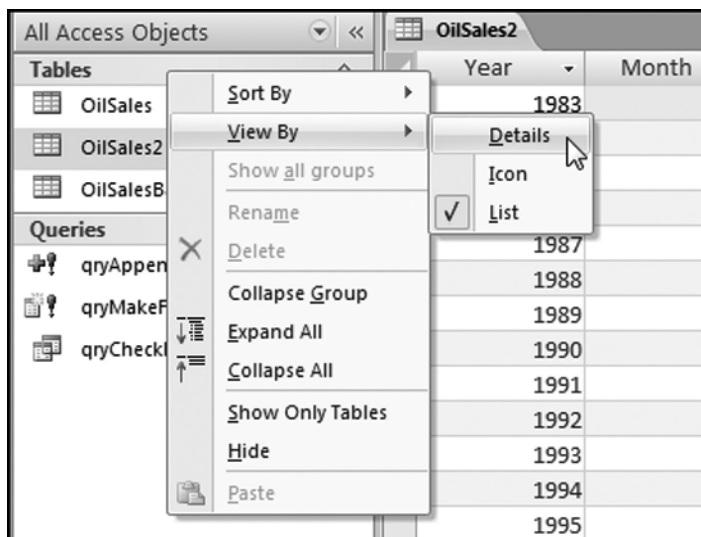


FIGURE 1.28 How to View the Access Object Details Instead of Just a List of the Objects

DATABASE LIMIT OF 2 GB

Access databases are limited to 2 gigabytes (2 GB). This limit applies to the database size when the database is saved and also while the database is being used. Assume that a database has one main table and the file size is 1.8 gigabytes. A query that selects most of the records would then double the size of the database to 3.6 gigabytes (the original table plus the query that selects most of the records). Access will not execute this query. The “size exceeded” error messages that Access displays do not clearly say “you have exceeded the Access maximum database size.” The user is simply supposed to realize why the query is not executing properly. There is a solution to the 2 GB limit in that the data can be housed in multiple databases (each say one gigabyte in size) and the main Access database can then be linked to those tables. Linked tables do not add to the size of a database. Users still have to be aware that if one links to two 1.5 GB tables and then runs two queries that each selects most of the records in each table, then the second of those two queries will not run if the first query is still open.

There are a few more solutions to the database size limit. In a forensic investigations project the database size can be kept down by deleting fields that are not needed. Deleting text fields is a big help with reducing the size of the database. Another option is to upgrade to a data analysis software program and IDEA (www.caseware.com) works well in an environment with large data sets.

MISCELLANEOUS ACCESS NOTES

This section reviews some miscellaneous aspects of Access that are relevant to forensic analytics. The list is based on personal experience and users should refer to one of the many comprehensive books on Access and/or the Microsoft website for more details.

- It is normal to format the output of a query for presentation purposes by (say) displaying only the significant digits of numbers. In Access all formatting must be done before the query is run. After the query is run the only formatting step that can be taken is to export the results to Excel and to format the results in Excel.
- Keep field names and table names short. Long field and table names cause extra work if these are used in calculated fields in queries.
- If a wrong field specification (e.g., define a date field as numeric) is entered when importing data into Access, then this issue cannot be corrected after the table has been created.
- It is good practice to create a backup copy of a table that will be queried frequently. Changes (deletions) to a table cannot be reversed. Access does not allow an “exit without saving” option.
- If a table is sorted immediately after creating the table then this sort is a “built-in” query that will run each time that the table is opened. This action can be undone.
- Dollar amounts should not be formatted as Double. The Currency data type should be used instead.

- It is a good practice to include a Primary key in tables. This optimizes computer performance for queries. The **Database Tools**→**Analyze**→**Analyze Performance** tool provides performance and other suggestions.
- To end a query that seems to have Access going in a loop use **Control + Break**.
- Access has a password encryption feature. The tool is accessed through **Database Tools**→**Database Tools**→**Encrypt with Password**. This tool is good enough to keep unsophisticated hackers at bay, but a determined tech-smart person could still work around the password.
- Including a switchboard with an application makes it look professional to another user and will allow users to create reports and to run queries with a single click. It is a good idea to practice creating a switchboard on a simple database with one or two tables and just a few queries. The Switchboard Manager helps with the process and it is accessed through **Database Tools**→**Database Tools**→**Switchboard Manager**.

Access has many Access Options. It is a good idea to use the “Compact on Close” option. This saves hard drive space and helps with the 2 GB limit. With large databases the Compact on Close procedure might take a few minutes.

SUMMARY

The chapter introduces Microsoft Access 2007 (Access) as a capable forensic analytics tool. Access is a Windows-based database program that keeps the tables, queries, and reports neatly compartmentalized. Access requires data to be housed in tables, calculations to be run as queries, and results to be shown in reports. Microsoft has a website for Access 2007, which contains excellent reference materials.

The usual forensic analytics starting point is to import the data into Access and to store it in a table. Tables are made up of fields, with each field storing one type of data for all the records. Records relate to one instance of a table subject, such as a book in a library’s collection of books. Field values are a single number, date, or text value relating to a record. A table with eight fields and 1,000 records would have 8,000 field values. The data import procedure is usually quite straightforward especially if the data is being imported from an Excel file or from another Access database.

Queries are the workhorses used in forensic analytics. Queries are used to perform calculations or to select records with specific attributes (e.g., all the transactions for vendor #2204). Queries are also used to append tables to each other, to create tables, to delete data, and to update the data in tables. Queries are also used to group data and to run calculations on the groups (e.g., sums and averages). Queries can also be used to identify duplicate or near-duplicate records.

It is usually necessary to do some data cleansing and some data reorganizing work on data downloaded from the Internet. This is because the way that data is accumulated in Excel worksheets does not work well in an Access database. This chapter shows how

26 ■ Using Access in Forensic Investigations

to use a series of append queries to create an Access table that is compatible with the required attributes of a table and the logic of Access queries.

The chapter reviews the need for adequate documentation in a forensic analytics setting, and Access has tools available for this purpose. The Database Documenter creates a complete record of the contents of the database. Access also allows forensic investigators to fully describe tables, queries, and reports in the database documentation. There is also a way to describe in detail each field in an Access table. Access also allows users to document the database at a high level in the Database Properties section. Documentation is important so that someone can understand the contents and queries in a database months or years down the road. The chapter concludes with notes related to formatting, field names, data-type specifications, passwords, and switchboards.